

Curating Large-scale Face Recognition Benchmark Test Sets

Dr. Nate Kalka, Brianna Maze

11/29/2018



For the best of reasons

Agenda

Disclaimer: This research is based upon work supported by the Office of the Director of National Intelligence (ODNI) and the Intelligence Advanced Research Projects Activity (IARPA). The views and conclusions contained herein are those of the authors and should not be interpreted as necessarily representing the official policies or endorsements, either expressed or implied, of ODNI, IARPA, or the U.S. Government. The U.S. Government is authorized to reproduce and distribute reprints for Governmental purposes notwithstanding any copyright annotation thereon.

Background



Dataset Curation Approaches

Crowd-Sourcing Approach

Hybrid Approach

Conclusion

Background: Levels of Difficulty in Face Recognition

	Frontal, Cooperative, Controlled	Near frontal, uncooperative	Full variation in A-PIE, uncooperative
Automated Face Detection Performance	Human	Near human	Limited
Automated FR Performance	Human	Near human	Limited



Background: Advances in State of the Art



COMPUTE RESOURCES

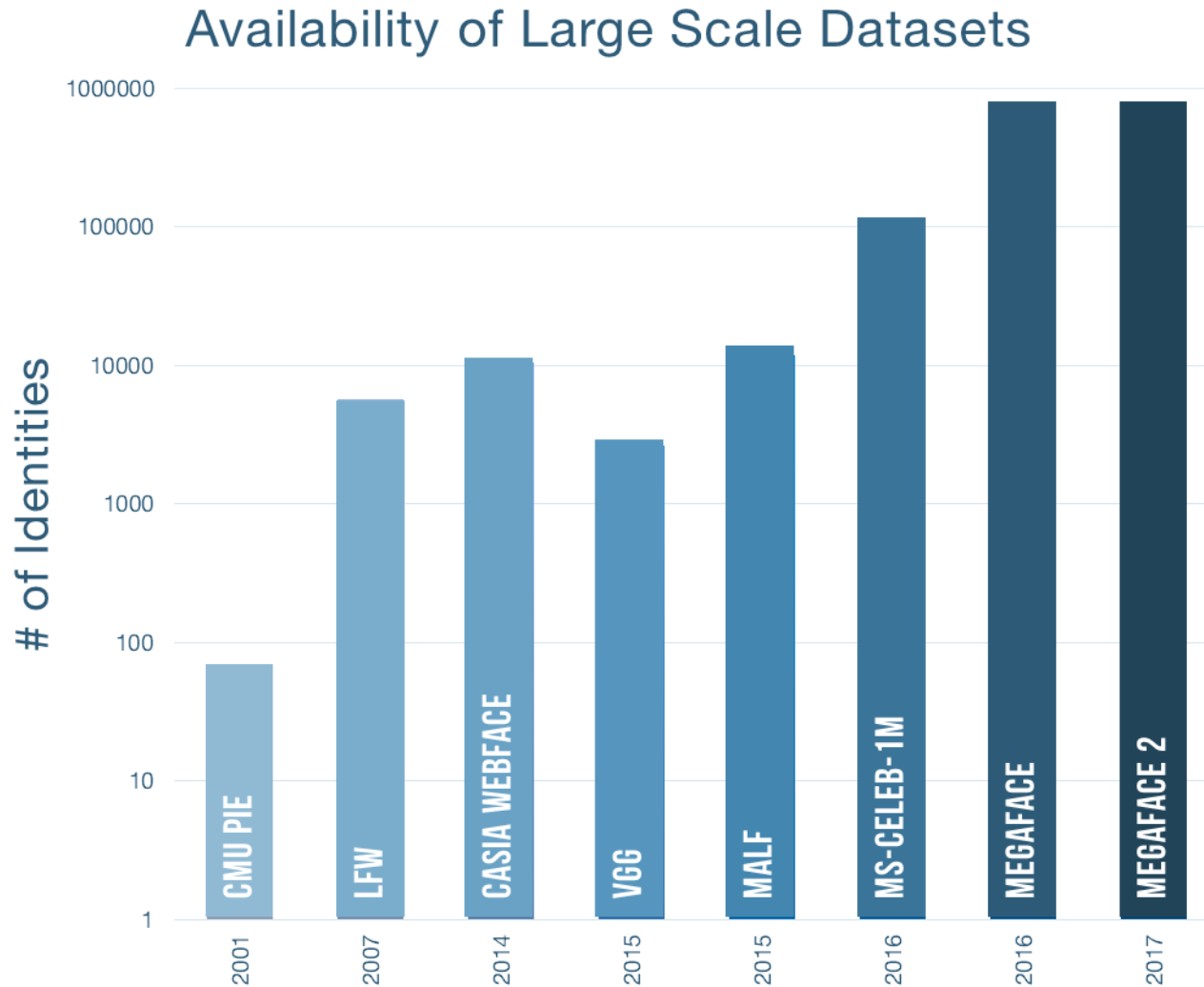


**DEEP LEARNING
ALGORITHMS**

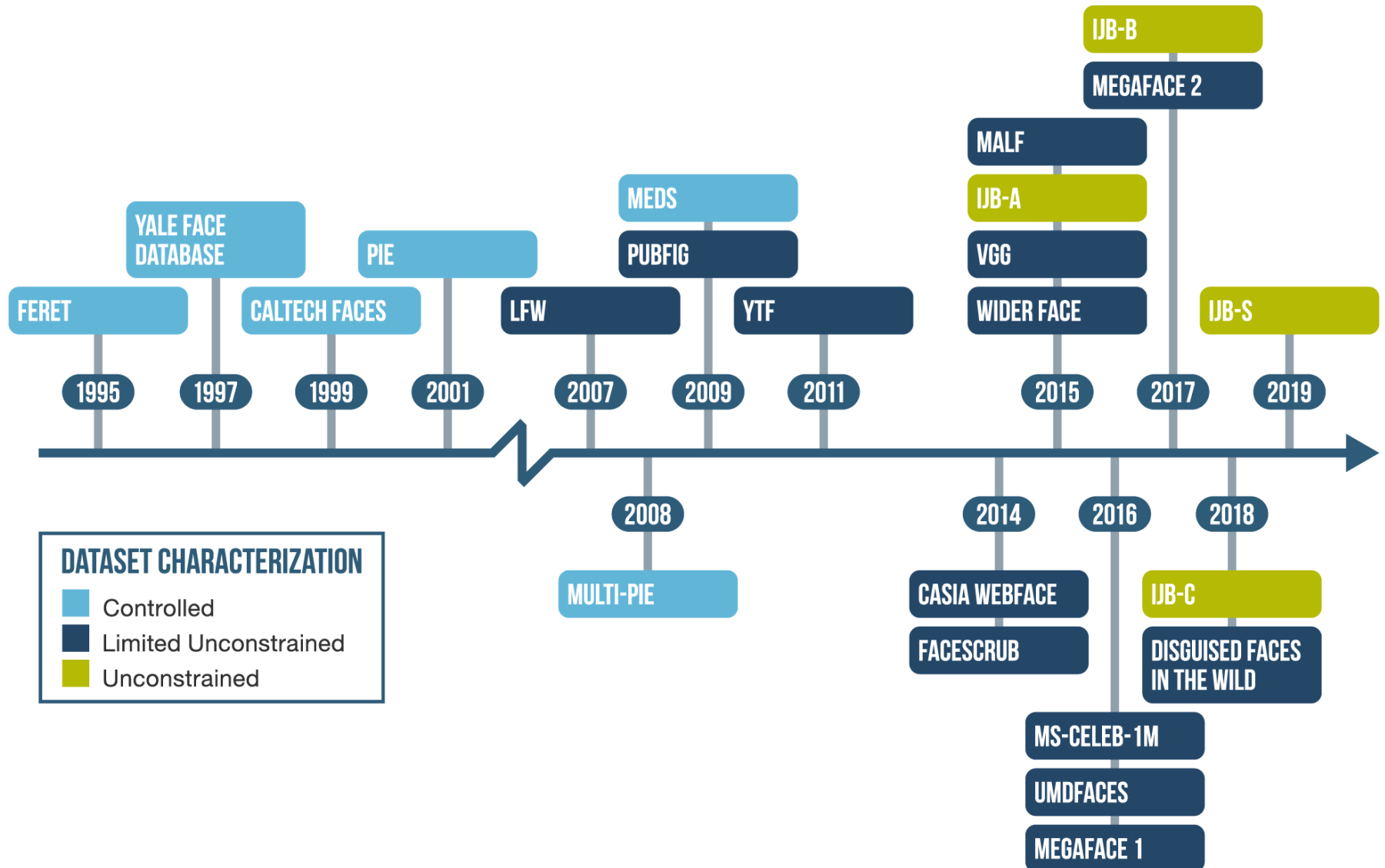


AVAILABILITY OF DATA

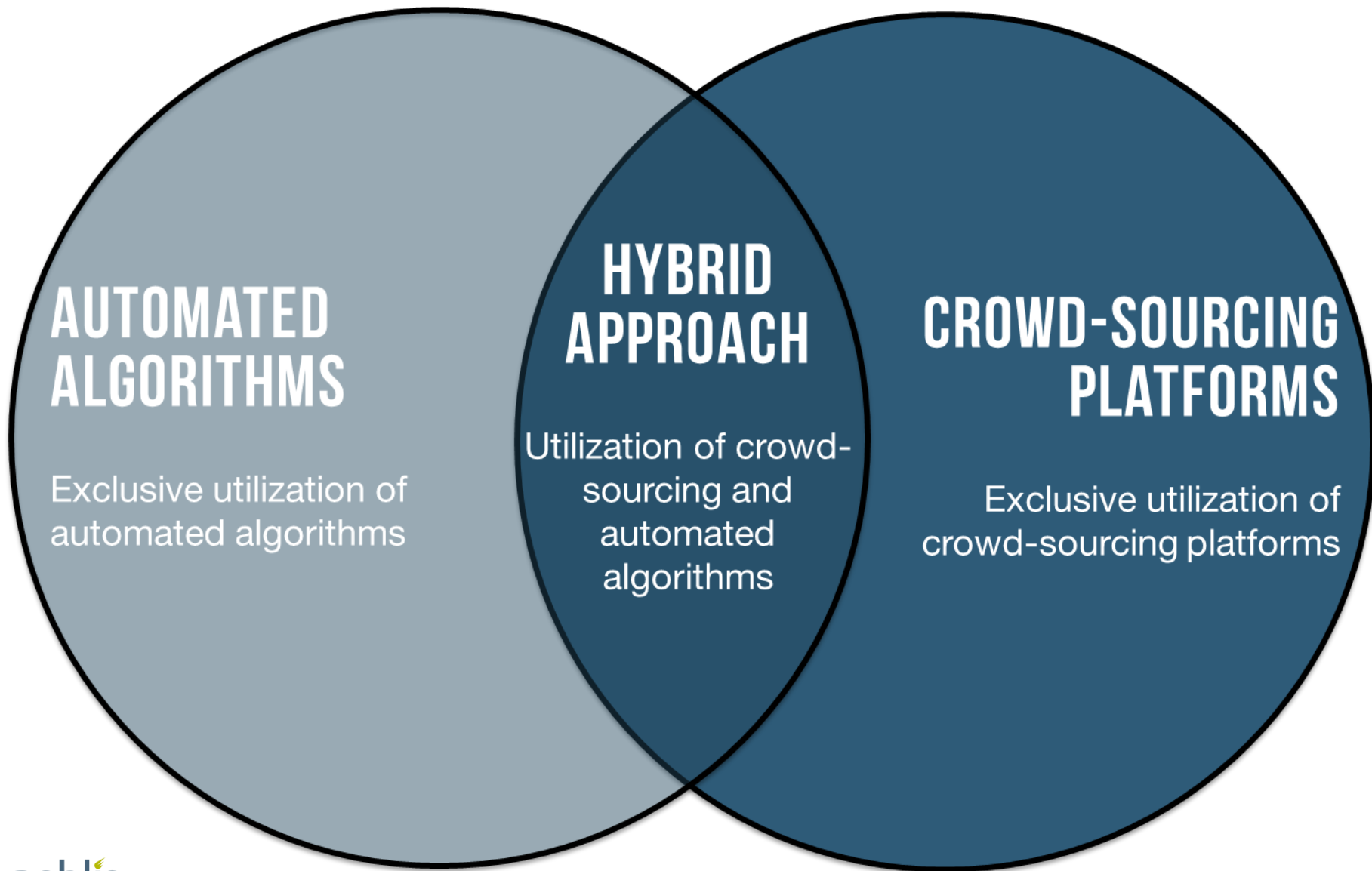
Background: Advances in State of the Art



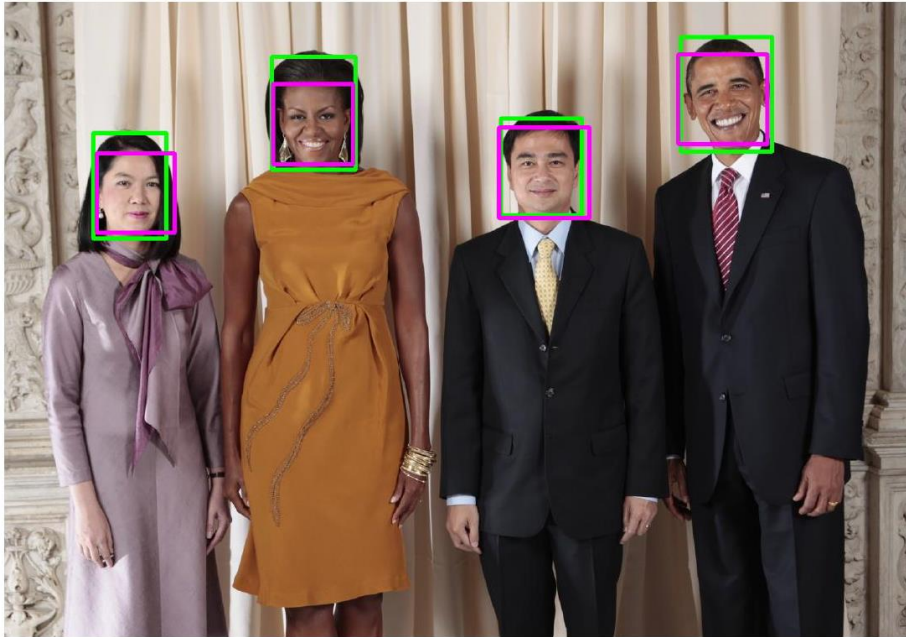
State of the Art Datasets



Dataset Curation Approaches



Crowd-Sourcing Approach



Crowd-Sourcing Approach



Crowd-Sourcing Annotations

Requester

Worker ID: A1A0CXMHTUN4E [🔗](#)

Hello, NoblisCV | [Sign Out](#)



HITs

[Dashboard](#)

[Qualifications](#)



[Filter](#)

All HITs

[Your HITs Queue](#)

HIT Groups (1-20 of 1108)

[+ Show Details](#)

[- Hide Details](#)

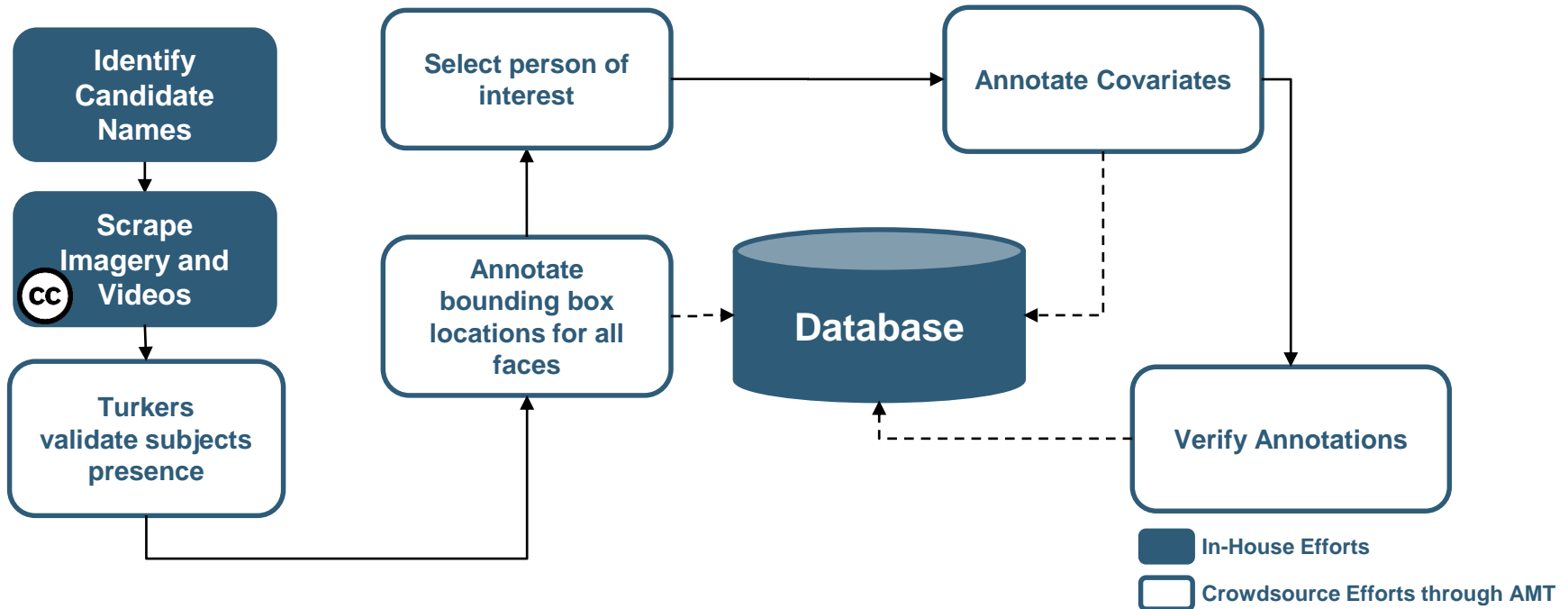
Items Per Page: **20** ▼

Requester	Title	HITS ▼	Reward ▼	Created ▼	Actions	
+ Victoria Reznyk	Which Product Type is the most relevant? (v3.1D)	23,252	\$0.05	18h ago	Preview	🔒 Qualify
+ mturk-sandbox	Categorize the purpose of information technology posts	10,023	\$0.02	5d ago	Preview	🔒 Qualify
+ ScoutIt	Classify Receipt	9,240	\$0.03	40s ago	Preview	🔒 Qualify
+ James Billings	Market Research Survey	6,630	\$0.01	19m ago	Preview	Accept & Work
+ Content Research	Clean Up How-To Questions	6,115	\$0.05	2h ago	Preview	🔒 Qualify

Human
Intelligence
Task (HIT)

HIT
Reward

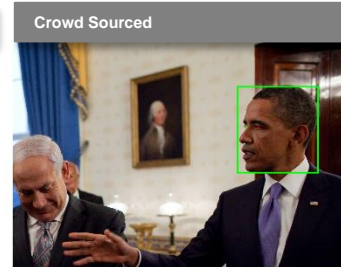
Crowd-Sourcing Approach



Start with subject imagery



Select BB for POI



Label image and subject attributes



Inspection by analyst

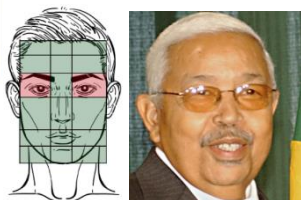
Annotations

CROWD-SOURCED

Face Detections



Occlusion Grid



Facial Hair

None

Beard

Mustache

Goatee



Environment

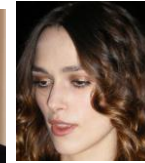
Indoor

Outdoor



Absolute Age

1 – 99



AUTOMATED/IN HOUSE

Skin Color

Cat. 1

Cat. 2

Cat. 3

Cat. 4

Cat. 5

Cat. 6



Pose θ (yaw)

$|\theta| < 15^\circ$

$15^\circ \leq |\theta| < 30^\circ$

$30^\circ \leq |\theta|$

Unknown



Gender

Male

Female



Geographic Region

North American

South America

Western Europe

South, West Africa

East Europe

East Africa, Middle East

Southeast Asia

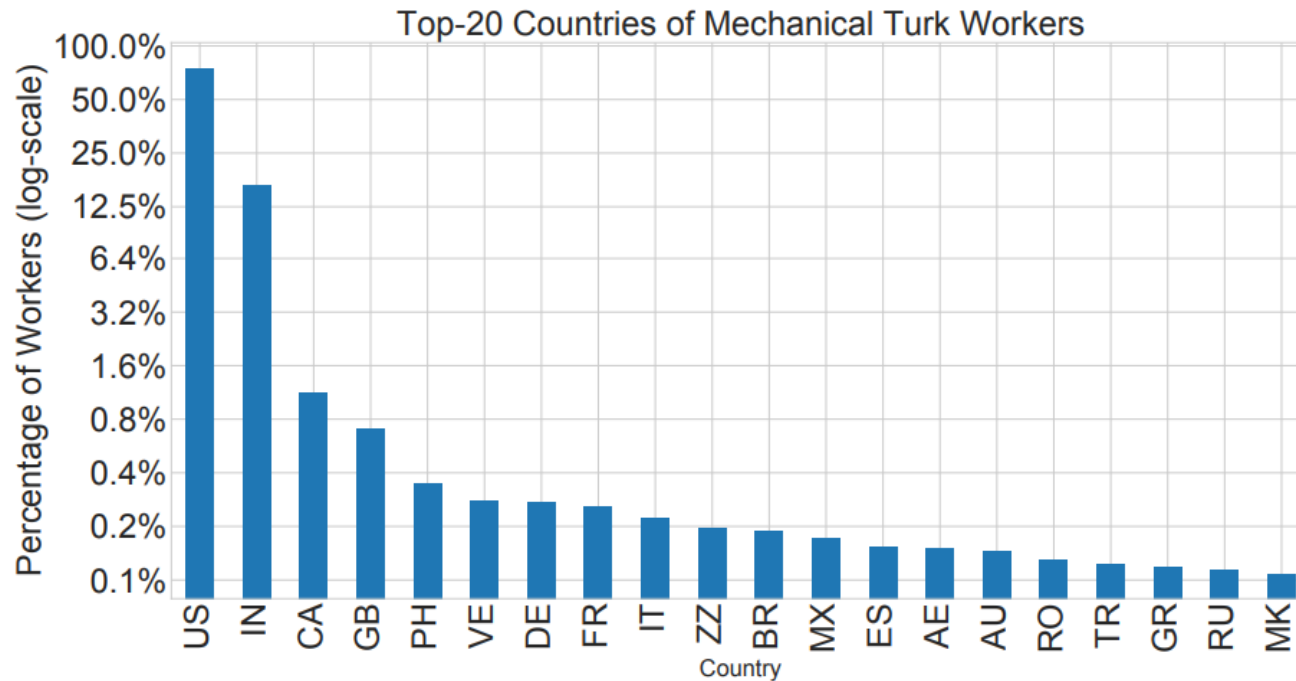
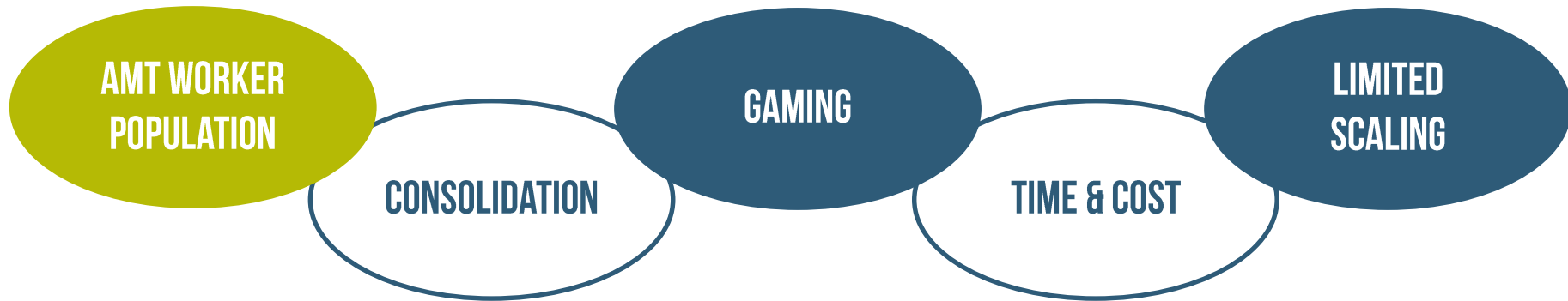
India

China

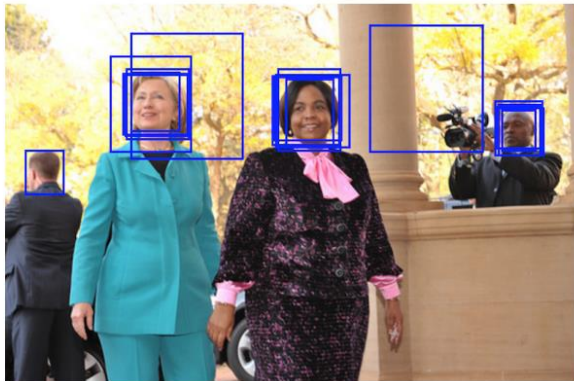
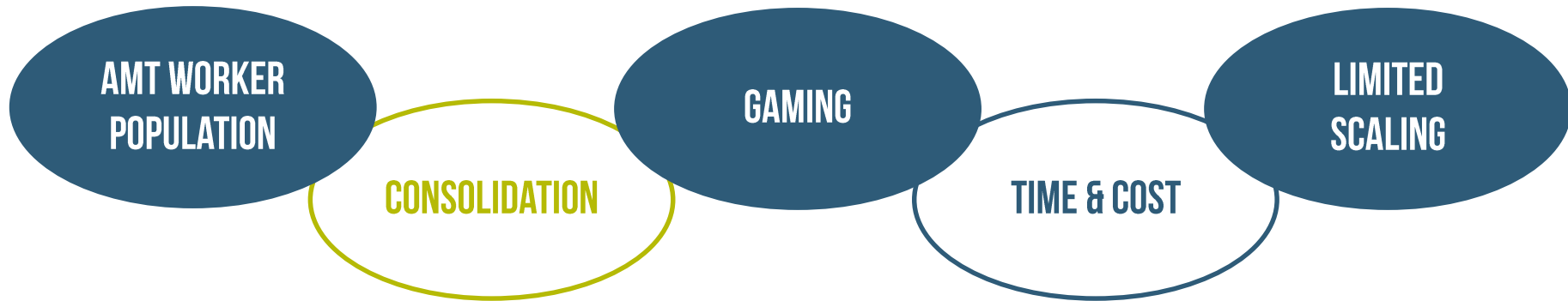
East Asia



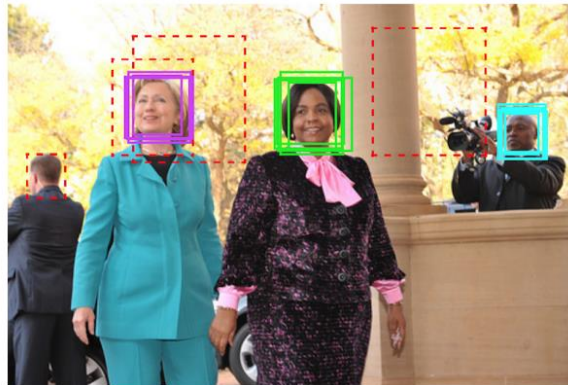
Limitations of Amazon Mechanical Turk



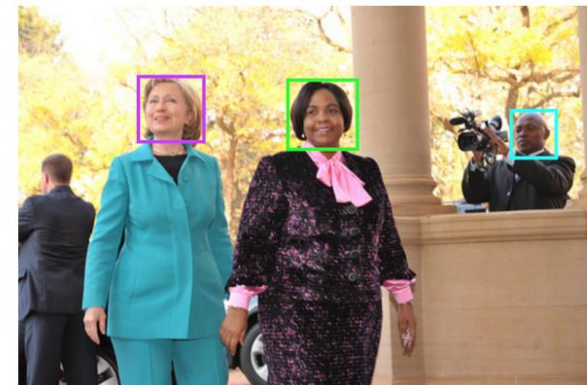
Limitations of Amazon Mechanical Turk



(a) Raw bounding box annotations.



(b) Clustered bounding box groups.



(c) Bounding boxes after consolidation.

Limitations of Amazon Mechanical Turk



Please disregard everything I said earlier—I lied to you and tried to play dumb, which I regret. I am also sorry for wasting your time.

When I received the rejections I panicked and I saw no other way but to lie to you. But after reading your reasonable reply, seeing that

This did not begin with malicious intent. The wage for the HIT was quite good, and I intended to work on it, so I tried hard to make it work on the keyboard. (You can implement the guides and keyboard bindings in your own code if you want to make your hits more convenient for the user.) But after a while I discovered that you can submit the HIT after completing only the first image.

The code is simple and I hope that you can use it to fix this exploit:

```
$(document).keypress(function(e){  
  if (e.which == 101){  
    $("#submitbutton").click();  
  }  
});
```

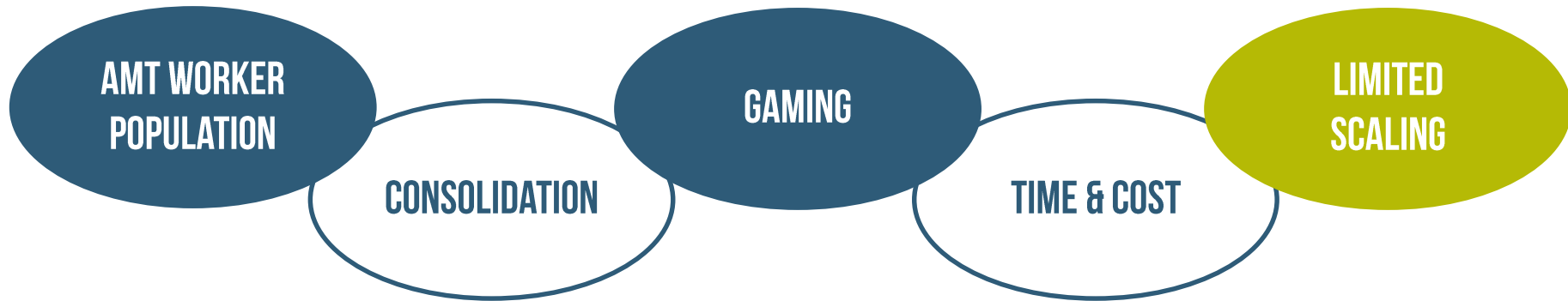
Limitations of Amazon Mechanical Turk



Ada... RE: Regarding Amazon Mechanical Turk HIT It's working now. I've done a good many of them. I guessed it just took time for the qualification to kick	Thu 9/28/2017 5:28 PM
Roc... RE: Regarding HIT Titled: Subject image review (WARNING: This HIT m... Thank you very much! You rock!!	Thu 9/28/2017 5:16 PM
Mit... RE: Regarding your Amazon Mechanical Turk HITs yes I requested it but was not auto granted	Thu 9/28/2017 4:52 PM
Roc... RE: Regarding HIT Titled: Subject image review (WARNING: This HIT m... Hi there! I checked and I still don't have it. I've requested again, but still waiting. I appreciate your help w	Thu 9/28/2017 4:45 PM
Jan... RE: Regarding Amazon Mechanical Turk HIT 35U0MRQMUJ2F3C7EN... Hi Lisa! Yes, that's perfect. As long as you can say it's the same person with certainty, then it's a good ima	Thu 9/28/2017 3:37 PM
Jenny Regarding Amazon Mechanical Turk HIT 3YZ7A3YHR54H2DGQ84MW... Message from Jenny (jftdesignstudios@hotmail.com)	Thu 9/28/2017 3:34 PM
lisa ... Regarding Amazon Mechanical Turk HIT 35U0MRQMUJ2F3C7EN0GF... Message from lisa mcdonald (mcdonalds16@gmail.com)	Thu 9/28/2017 3:33 PM
Jan... RE: Regarding Amazon Mechanical Turk HIT 359AP8GAGGVGI05DZBB... Hi Hannah, Thanks for reaching out to us about your mistake. In the future, just be more careful, but sin	Thu 9/28/2017 3:28 PM
Han... Regarding Amazon Mechanical Turk HIT 359AP8GAGGVGI05DZBBTF1S... Message from Hannah K. McCoy (writer.hk1432@gmail.com)	Thu 9/28/2017 3:25 PM
Nic... RE: Regarding your Amazon Mechanical Turk HITs Thank you!	Thu 9/28/2017 3:13 PM
Jan... RE: Regarding your Amazon Mechanical Turk HITs Hi Nick, Thanks for your patience. I manually granted you the qual.	Thu 9/28/2017 3:11 PM
Nic... RE: Regarding your Amazon Mechanical Turk HITs I requested it -- I believe a few hours ago -- but it hasn't yet been granted. Is there usually a notable dela	Thu 9/28/2017 3:07 PM
Jan... RE: Regarding HIT Titled: Subject image review (WARNING: This HIT m... Hi Rochelle, You should be able to click "Request Qualification" and it should automatically grant you the	Thu 9/28/2017 3:04 PM
Jan... RE: Regarding Amazon Mechanical Turk HIT 3UXQ63NLAAXWNG29P5... Hi Pedro,	Thu 9/28/2017 3:03 PM
Roc... Regarding HIT Titled: Subject image review (WARNING: This HIT may c... Message from Rochelle Nelson (rochelle@trevarrow@gmail.com)	Thu 9/28/2017 3:02 PM
Ped... Regarding Amazon Mechanical Turk HIT 3UXQ63NLAAXWNG29P5G8N... Message from Pedro L. Solis (beachwavesentropy1@outlook.com)	Thu 9/28/2017 3:01 PM

All Reviews	Flagged By You	Your Reviews	Order by edit date
AMT Requester	Rating [info]	Description	
HITS A1A0CXMHXTUN4E Averages » HIT Group » Review Requester »	FAIR: 5 / 5 FAST: 5 / 5 PAY: 1 / 5 COMM: NO DATA HIT APPROVED	HIT(s): Duplicate subject review This requester has posted different types of HITs in the past. Different HITs pay differently. Some were OK, but some were not. This one in particular, you have to look at lots of images scrolling down to match images. I found myself spending about closed to 5 minutes or more to submit a HIT, which is not really worthwhile. I would not recommend putting this on PANDA since there is time limit of 20 minutes. Like I said, there are also lots of images to go through scrolling down and clicking Next. Feb 06 2018 Gturk	
NoblisCV A1A0CXMHXTUN4E Averages » HIT Group » Review Requester »	FAIR: 5 / 5 FAST: 5 / 5 PAY: 5 / 5 COMM: 5 / 5 HIT APPROVED	HIT(s): Occlusion Grid Very fair with amount paid and the time needed for completing this hits. If you put a honest effort and read his instructions its very easy to comprehend. I submitted 10 hits worth \$0.50 each, each hit took about 3 to 4 minutes to complete and approved with in 3 days. Mar 11 2017 Honest Mturk Amazonia	
NoblisCV A1A0CXMHXTUN4E Averages » HIT Group » Review Requester »	FAIR: 5 / 5 FAST: 4 / 5 PAY: 4 / 5 COMM: 4 / 5 HIT APPROVED	I have had success with them. I have never had a rejection and they have responded to some of my questions. I have done a lot of hits and can't complain about them. Mar 08 2017 Go Blue	

Limitations of Amazon Mechanical Turk



NEXT ECONOMY

The Internet Is Enabling a New Kind of Poorly Paid Hell

For some Americans, sub-minimum-wage online tasks are the only work available.

ALANA SEMUELS JAN 23, 2018

AMAZON'S TURKER CROWD HAS HAD ENOUGH

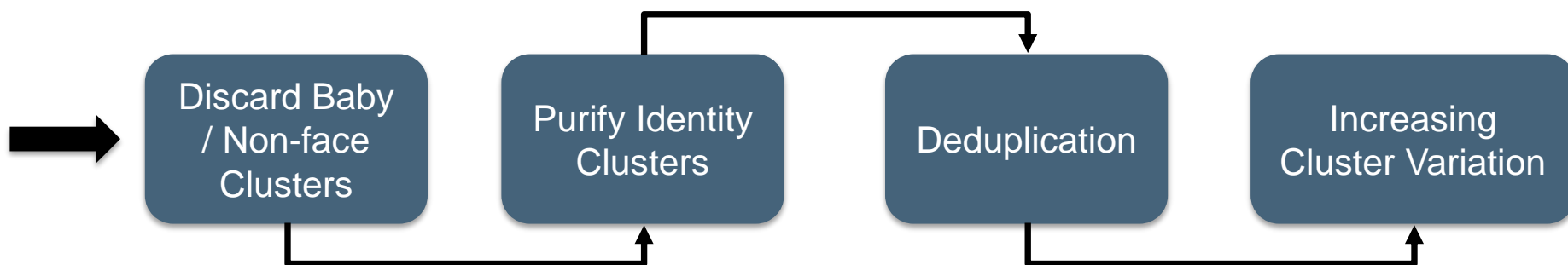
EMILY DREYFUSS SECURITY 08.17.18 11:38 AM

A BOT PANIC HITS AMAZON'S MECHANICAL TURK

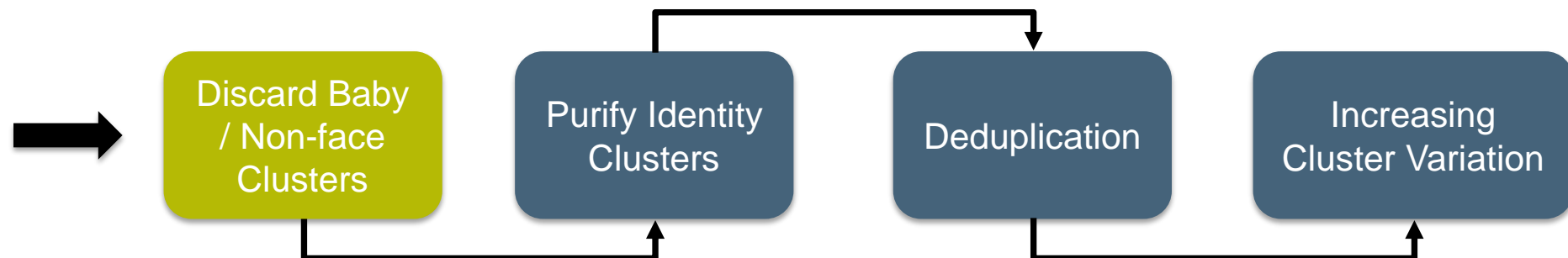
Hybrid Approach

Curating a Large Scale Testing Set

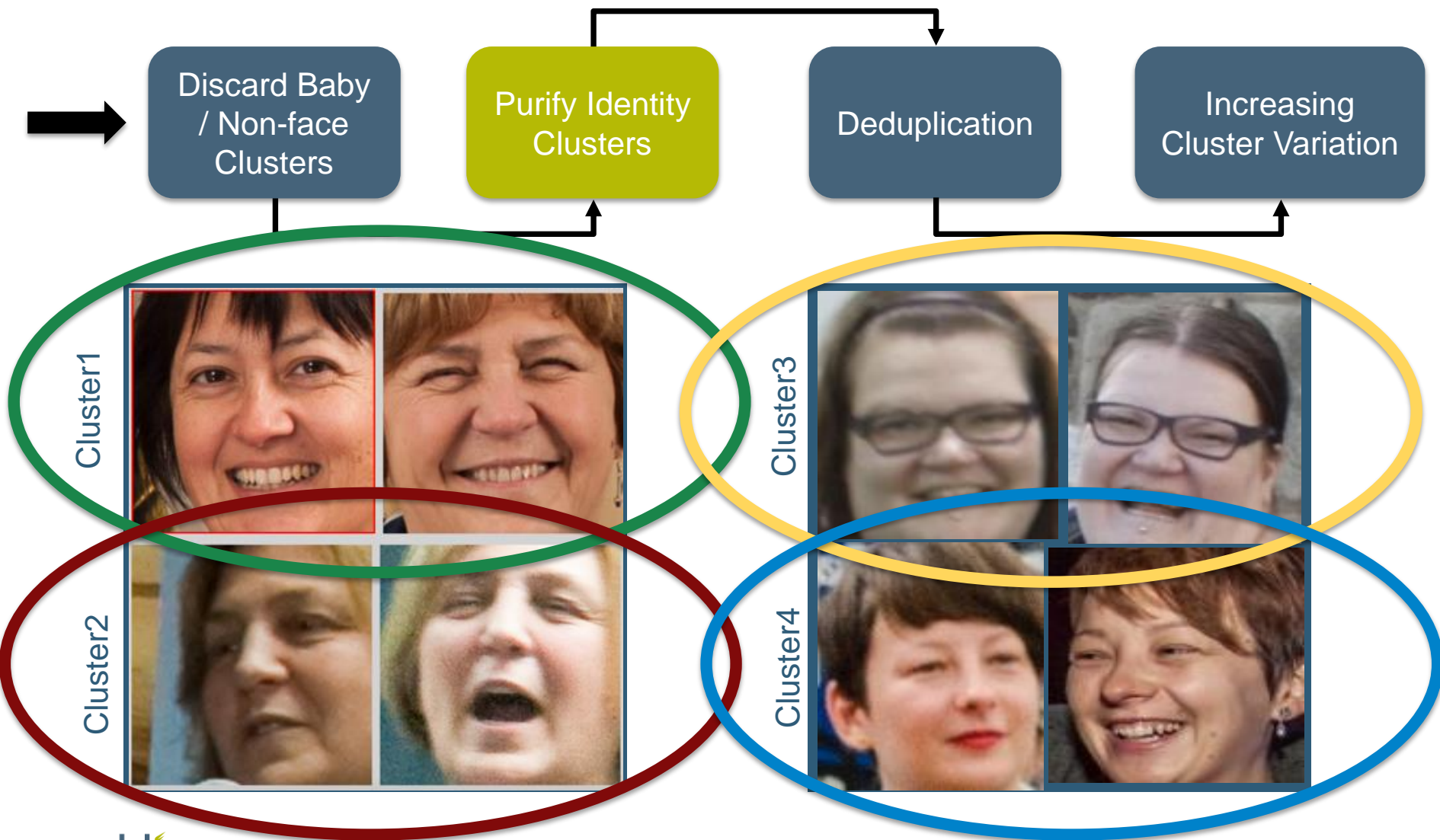
- **4.7 Million** photos
- **672,057** Unique Identities
- **7** Mean photos / person
- Subset of the Yahoo Flickr Creative Commons 100 Million set
- Intended for use as a training set
- How do we transform into a useful test evaluation set?



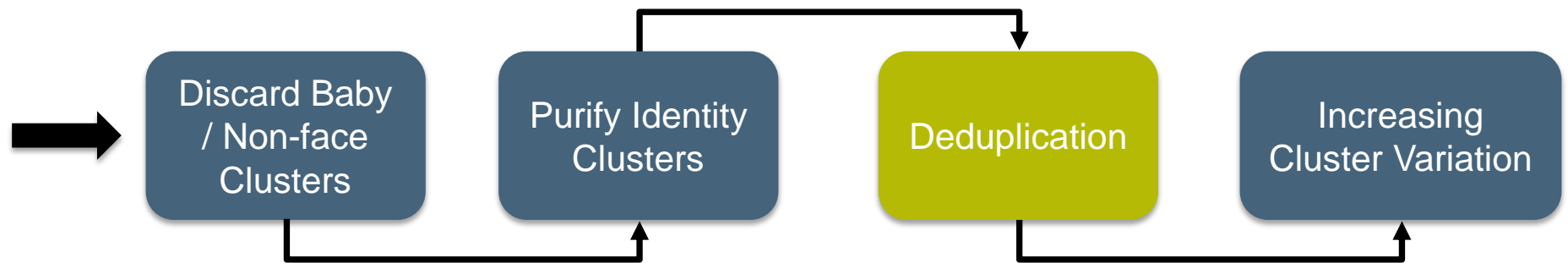
Discard non-face / baby clusters



Purify Identity Clusters



Deduplication



MERGE



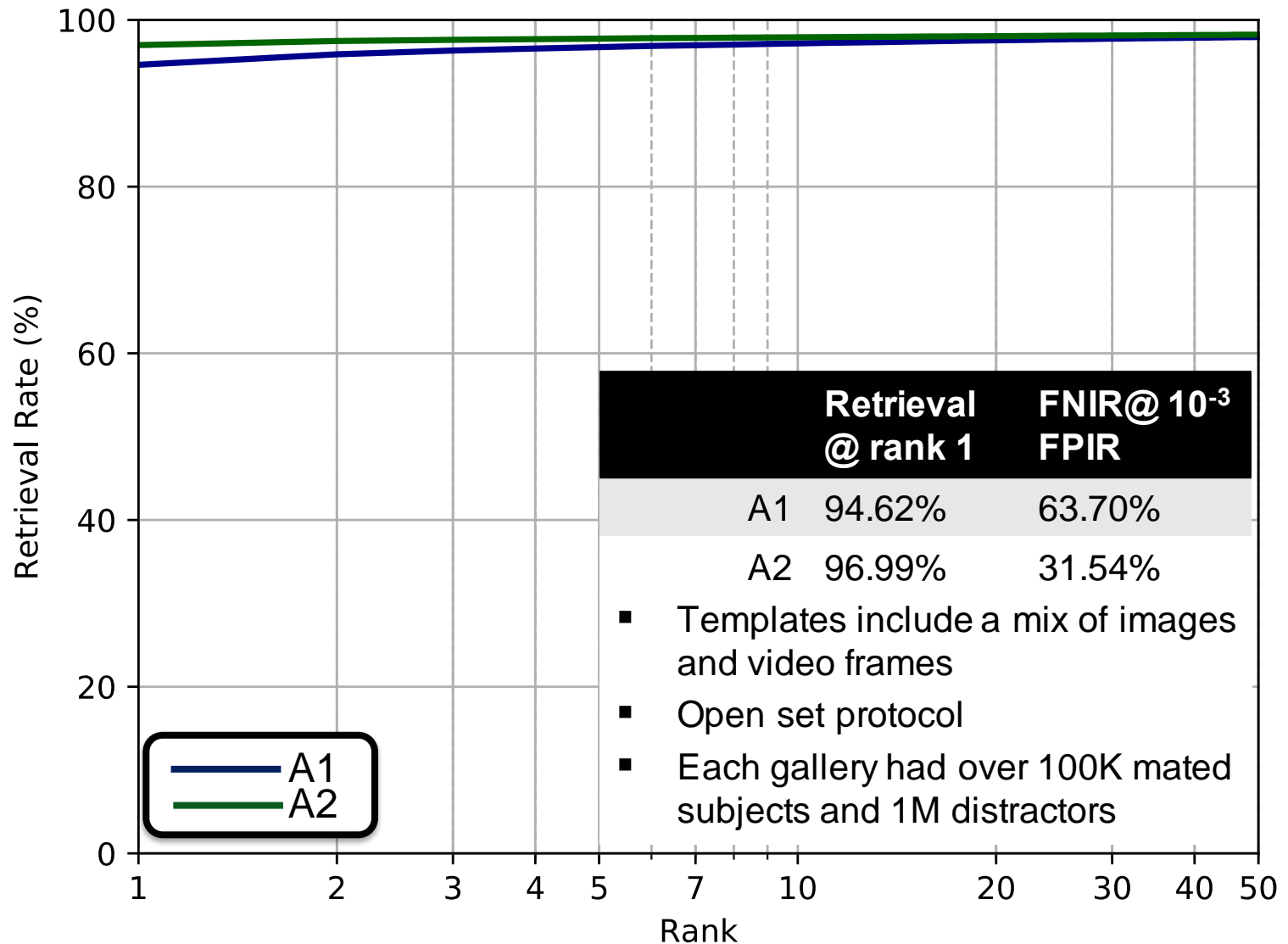
Final Numbers

	# Images	# Identity Clusters	# images / cluster
Original	4.7M	672,057	7
Cleaned	2.9M	471,231	6.26

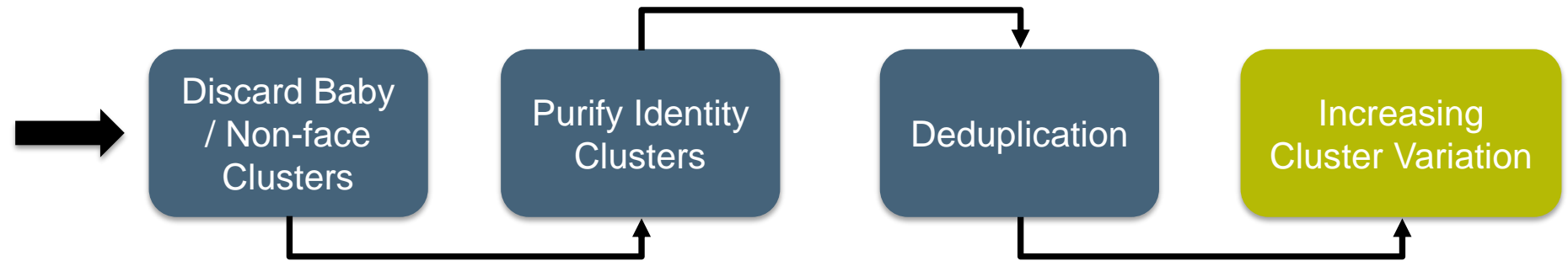
- Over 2.4M annotations
- Protocols for 1:1, closed and open set 1:N Identification, and training
- Discarded 61,216 baby clusters
- Discarded 98,407 non-face / chimeric clusters
- Removed 5,840 clusters with expired media
- Merged 35,363 clusters



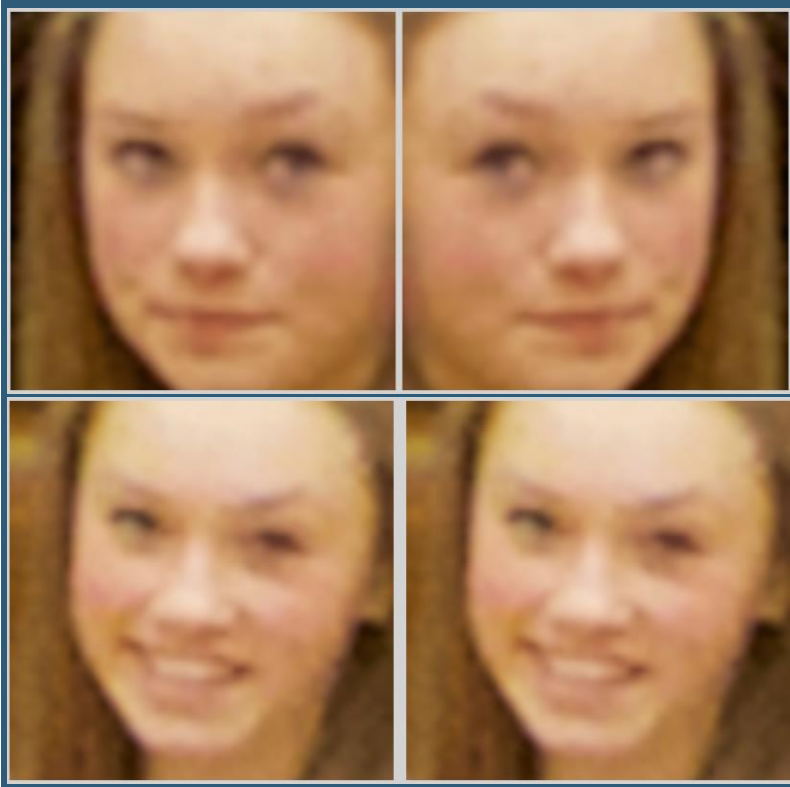
1:N Identification Performance



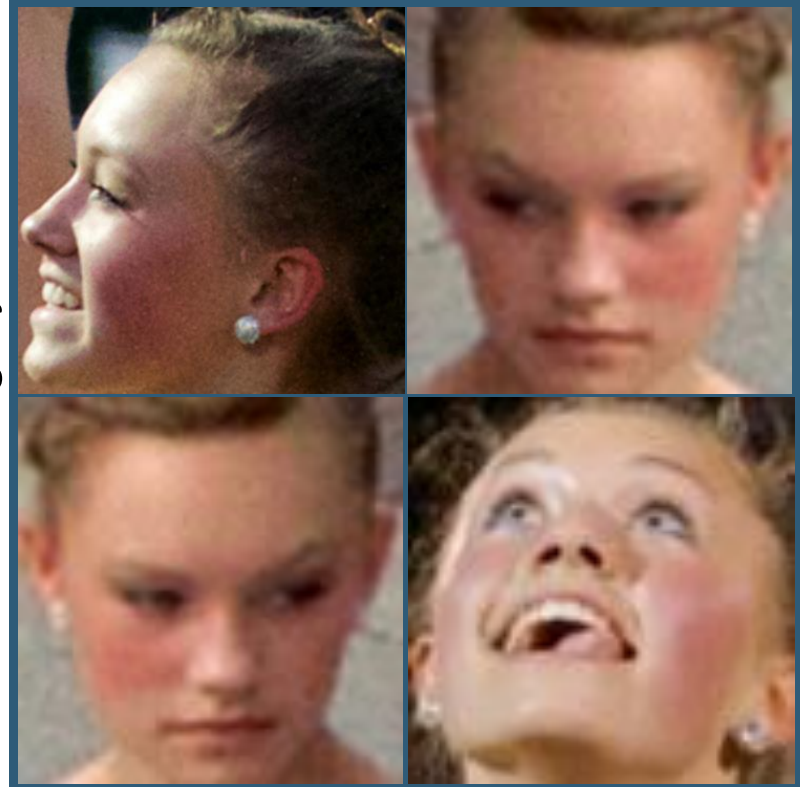
Increase Cluster Variation



Clustered automatically



Found among same user's imagery



Questions?



Nathan.Kalka@Noblis.org



Brianna.Maze@Noblis.org